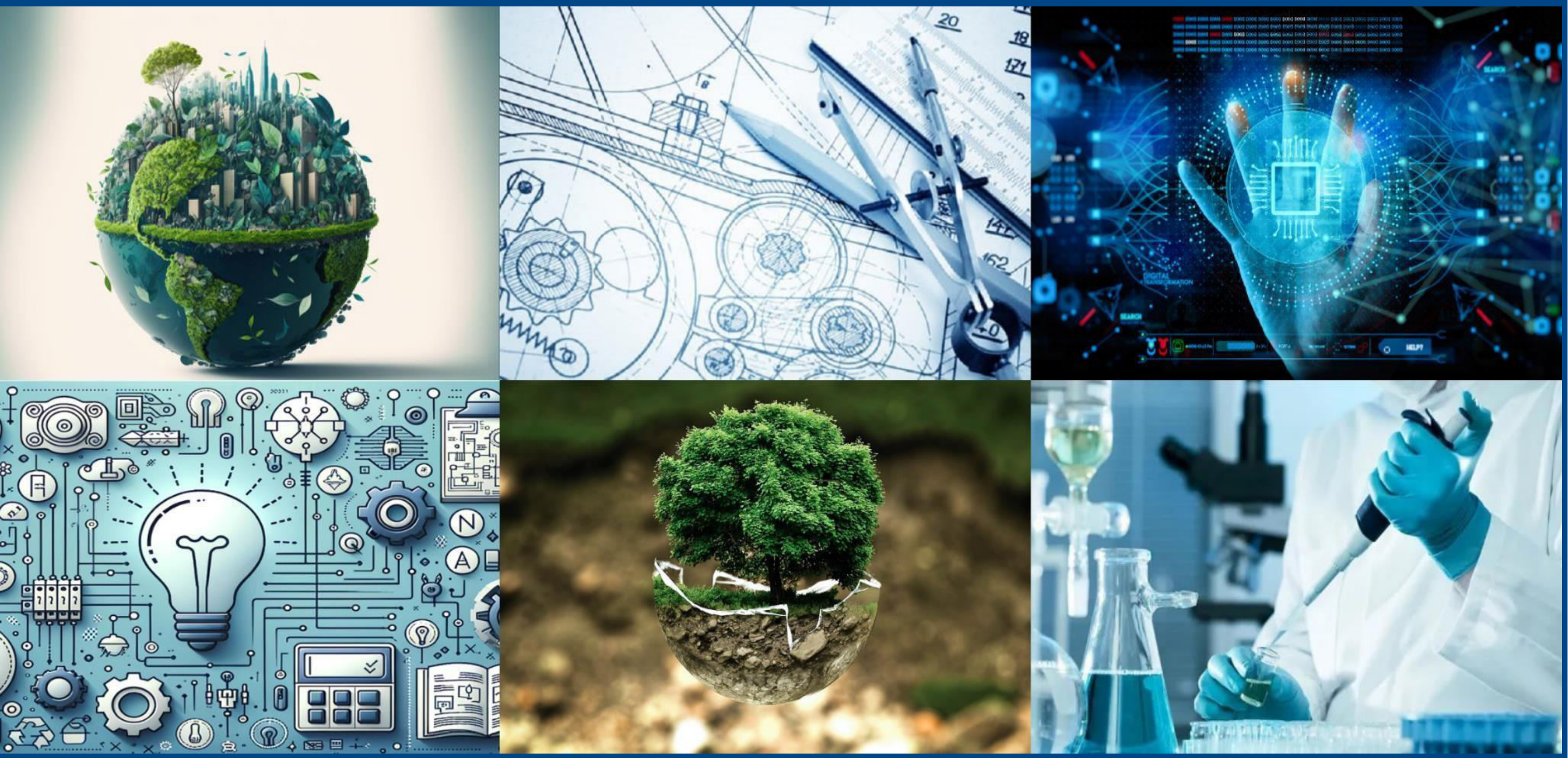
ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 8, Issue 12, December 2025

www.ijmrset.com | ijmrset@gmail.com | +91 6381 907 438 | +91 63819 07438

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Advanced Movie Recommendation Engine using Data Science

Sneha H Iti[1], Pooja Taragar[2]

PG Student, Dept. of MCA, City Engineering College, Bengaluru, India[1]

Assistant Professor, Dept. of MCA, City Engineering College, Bengaluru, India[2]

ABSTRACT: Movie recommendation systems are important for streaming services because they assist individuals in finding movies they like. This project recommends films based on specifics such as actors, directors, genres, and descriptions. It employs a content-based approach that examines movie text data and uses cosine-similarity and TF-IDF to identify commonalities. The most similar film to one a user enjoys are suggested by the system. This provides a better, more customized experience and speeds up the movie-finding process. How well the algorithm recommends movies to consumers is a key indicator of its effectiveness.

KEYWORDS: Movie recommendation, content-based filtering, TF-IDF, cosine similarity, machine learning, Streamlit**.**

I. INTRODUCTION

With so many films available on streaming services, It's getting harder for you to find movies that suit their tastes. By making movie recommendations based on previous user activity, movie specifics, or popular trends, recommendation systems assist in resolving this issue. Common techniques include collaborative filtering, which relies on user activity, and content-based filtering, which examines movie features. In this study, movie details like descriptions, genres, directors, and actors are analyzed using a content-based method with TF-IDF. The method finds and movies that are most comparable to a chosen one by comparing them using cosine similarity. Because it provides tailored recommendations without requiring a lot of user information, this is effective for both new and returning customers.

Aims: By providing personalized movie recommendations, reducing search time, and improving streaming platform content discovery, we hope to improve the user experience. The goal is to Establish a system that makes movie recommendations based on content for similar films depending on textual attributes such as descriptions, genres, directors, and actors using TF-IDF vectorization and cosine similarity.

II. LITERATURE SURVEY

1. To enhance movie suggestions, Sarwar et al. (2001) presented an item-based collaborative filtering technique. This method makes the system faster and more scalable by comparing movies rather than users.
2. Matrix factorization algorithms that uncover latent correlations between users and movies were proposed by Koren, Bell, and Volinsky (2009). This technique helped lower prediction mistakes like RMSE and gained a lot of popularity during the Netflix Prize competition.
3. Lops, de Gemmis, and Semeraro (2011) investigated content-based recommendation algorithms that utilize of movie attributes including descriptions, cast, and genre. According to their findings, this strategy offers more sensible suggestions for new users, but it occasionally makes overly similar movie suggestions
4. Burke (2017) examined hybrid recommendation systems, which blend content-based and collaborative filtering methods. As stated by hybrid systems offer more precise recommendations while resolving issues like cold-start and data sparsity.
5. Deep learning-based recommendation systems were examined by Zhang et al. (2019). The study described how, in comparison to conventional techniques, neural networks and autoencoders are better able to comprehend intricate relationships between users and movies.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

6. By examining user evaluations, Tripathy, Anand, and Rath (2020) incorporated sentiment analysisinto movie recommendation systems. This made it easier for the system to comprehend user feedback and provide more pertinent movie recommendations.
7. Context-aware recommendation algorithms that take into account elements such as time, location, and user mood were introduced by Adomavicius and Tuzhilin (2021). Making use of this data enhances the caliber and customization of suggestions.
8. Wu et al. (2022) modeled user and movie interactions as graphs using graph neural networks. Recommendation performance was enhanced by this method, particularly for sparse and complex datasets.

III. METHODOLOGY

Gathering movie data from online resources, including movie databases, is the first stage of the Advanced Movie Recommendation System. Movie titles, genres, user ratings, and reviews are all included in this. The system uses this data to better comprehend customer preferences and movies.

The data is cleansed and arranged after it has been gathered. Any missing or superfluous data is eliminated, and ratings are re transformed into a standard. Genres and other movie details are transformed into a format that the system can comprehend. Additionally, user reviews are cleansed to identify insightful viewpoints.

The system then examines the data to comprehend user behavior. It looks at user ratings, popular genres, and movies that are seen regularly. This aids the system in identifying trends and comprehending user preferences.

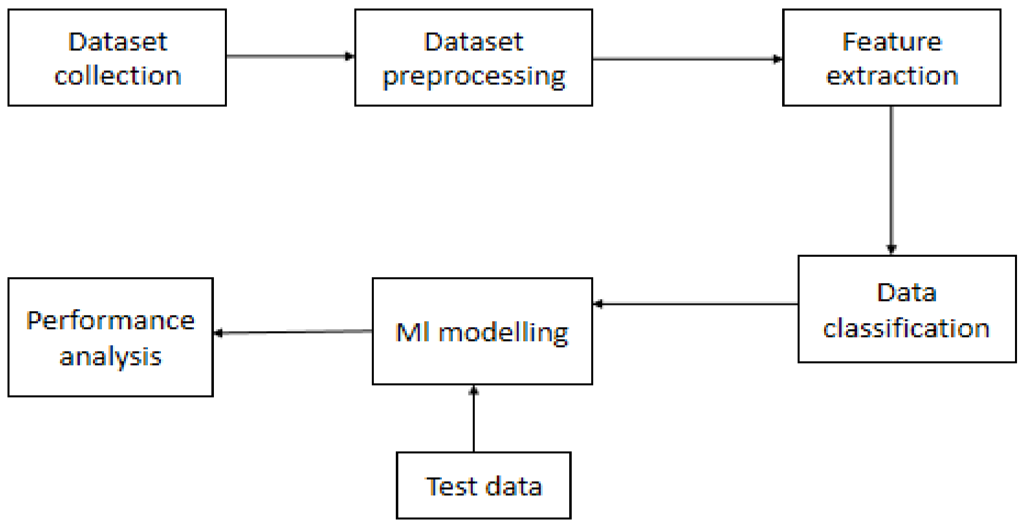


Figure 1: Methodology

IV. SYSTEM DESIGN

Data science and machine-learning are used by a sophisticated movie recommendation system to comprehend user preferences and make movie recommendations based on those preferences.

The system users are able to register, log in, and look for films, rate them, and watch suggested films. The system gathers user activity, including searches, clicks, ratings, and movies viewed. Additionally, it keeps movie information in a database, including the genre, cast, director, and description.

© 2025 IJMRSET | Volume 8, Issue 12, December 2025| DOI:10.15680/IJMRSET.2025.0812031

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The main part of the system is the recommendation engine. By examining the collected data, it recommends movies. It uses collaborative filtering to locate people who have similar interests as you content-based screening to suggest movies that are similar to those the user liked, and a hybrid approach that combines the two to increase accuracy. The machine-learning component examines the data, searches for patterns, and builds models to score the recommended movies. Over time, the system improves by using input to learn from new user activities.

The system is designed to safely handle a lot of users and movies.Python and Flask or Django are used to build the back-end, which facilitates the development of API s and real-time communication.

V. SYSTEM ARCHITECTURE & DESIGN

The data processing layer prepares user and movie data for training and prediction to enable this procedure. It chooses key features, cleans the data, and transforms it into a format that may be used. All data, including user profiles, movie information, ratings, and previous suggestions is stored in the database layer. This facilitates fast access when needed and helps maintain the accuracy of the data.

To obtain better and more comprehensive movie information, the system can additionally make use of outside resources like Movie-Lens. The sophisticated movie recommendation system is dependable and practical in real-world applications because to this well-organized and transparent design, which also enhances system speed, accuracy, and user experience.

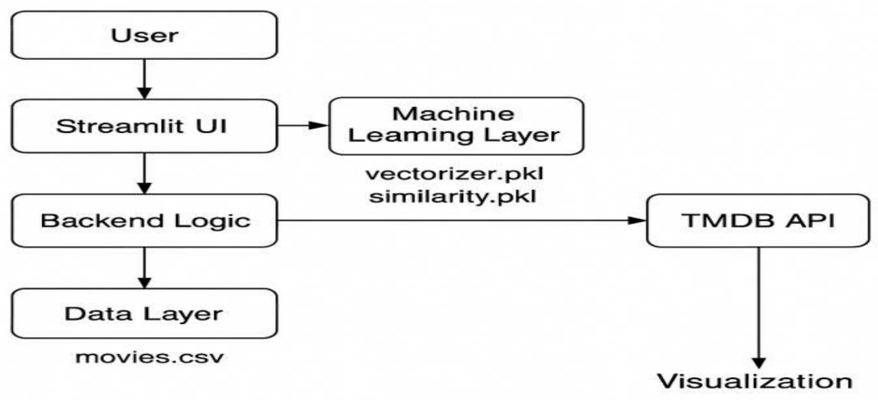


Figure 2: System Architecture

VI. IMPLEMENTATION

Movie information and reviews from users are gathered from trustworthy sources like Movie-Lens or TMDB in order to create a sophisticated movie recommendation system. User IDs, movie IDs, ratings, genres, and descriptions are all included in this data.

First, duplicate entries are eliminated, missing values are fixed, and ratings are adjusted to a common scale. Text and category Data is converted to numerical values that the system can comprehend using methods like TF-IDF and one-hot encoding. In order to represent movies according to their content and user interactions, key features are then extracted.

The system employs a number of recommendation techniques, including content-based filtering, which uses cosine similarity to recommend movies that person has a prior liked, and collaborative filtering, which forecasts user preferences depending on how comparable users or objects behave or items using matrix factorization methods such as Singular Value Decomposition (SVD). To improve accuracy and overcome limitations such as the cold-start problem, content-based and collaborative approaches are combined to develop a hybrid recommendation model.

The trained model provides a ranked list of movies to the front-end interface, which provides real-time recommendations. User feedback in the format of ratings and interactions is continuously recorded in order to retrain or update the model and improve accuracy over time. The system is then assessed for accuracy, scalability, and performance before being deployed on a cloud platform to enable real-world application

VII. RESULTS & DISCUSSION

The findings and evaluation of the advanced movie recommendation system demonstrate how effectively machine-learning and data science techniques can be applied to produce accurate and customized movie suggestions. After deployment, popular performance measurements such as Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) were used to evaluate the accuracy of projected ratings.

The experiment's Findings indicated that the hybrid approach produced more relevant recommendations and lower RMSE values than both collaborative and individual content-based filtering techniques. User-based collaborative filtering worked well for users with sufficient interaction history, but content-based filtering was successful in resolving the cold-start problem for new users. Combining the two approaches improved the system's overall robustness and suggestion quality.

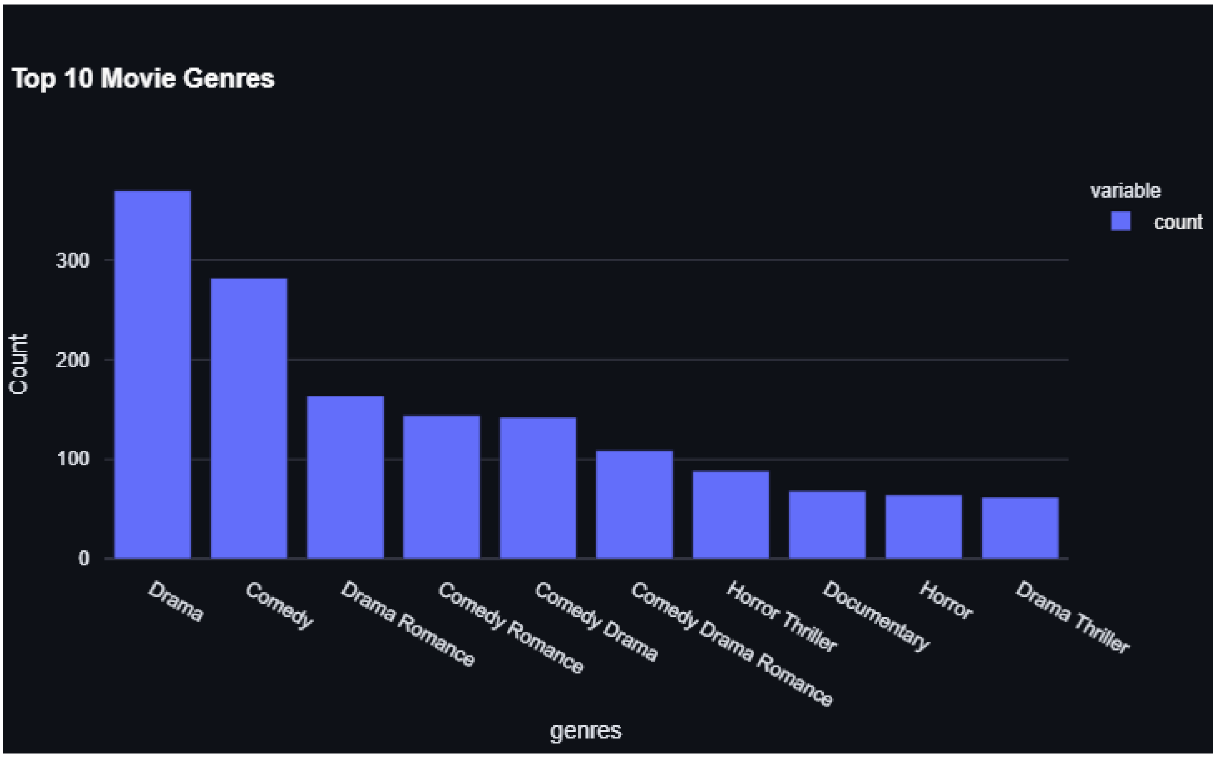


Figure 3: Dataset Visualization

VIII. CONCLUSION

The Conclusion of movie recommendation system demonstrates how machine-learning and data science can be used to propose films that suit a user's preferences. The system learns users' tastes by examining what they watch, like, and search for. This data is then used by machine-learning models to suggest films that the user is more likely to appreciate. By combining content-based filtering, collaborative filtering and a hybrid method, the system effectively addresses the drawbacks of traditional recommendation systems, such as cold-start issues and poor suggestion accuracy. This makes finding good movies faster, easier, and more personalized than traditional methods.The integration of user feedback and continuous learning allows the system to adapt to changing user preferences over time. Results from experiments

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

indicate better forecasting accuracy and higher user satisfaction when compared to conventional methods. Despite persistent problems with scalability and data sparsity, the system architecture and implementation demonstrate excellent potential for real-world application.

REFERENCES

1) Shapira, B., Ricci, F., and Rokach, L. (2011). Handbook of Recommender Systems. Springer.
Adomavicius, G., and Tuzhilin, A. (2005). A review of the of the most advanced and potential extensions for the upcoming generation of recommender systems.IEEE information and Data Transaction and Data Engineering, 17(6), 734-749.
2) Sarwar, B., Riedl, J., Konstan, J., and Karypis, G. (2001). recommendation methods for cooperative filtering that are item-based. The 10th International Conference proceedings World-Wide-Web Conference (WWW).
3) Shapira, B., Rokach, L., and Ricci, F. (2015). Recommender Systems: Overview and Difficulties. Springer.
4) C. C. Aggarwal (2016). The textbook on recommend systems. Springer.
5) Y. Koren, R. Bell & C. Volinsky (2009) — "Matrix Factorization Techniques for Recommender Systems" — seminal survey on MF and the Netflix Prize approaches
6) Slope One (Lemire & Maclachlan, 2005) — simple yet effective collaborative filtering predictors often used as baselines.
7) Li, Y., Chen, T., & Li, Z. (2022)- A Combination Movie Suggestion System Employing Deep Learning IEEE Access.Combines content features with collaborative filtering.

INTERNATIONAL JOURNAL OF
MULTIDISCIPLINARY RESEARCH
IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com